

SEQUENCE ALIGNMENT

FT	VARSPLOC	1143	1203	WGASSSLPQSPARVTTSMVTVAHPPLPLPGAYIHAPAPPEPW SPANTSSGNSSRGPGATG -> PEEI (in isoform 2).
FT				/FTID=VSP_004542.
FT				E -> Q.
FT	VARIANT	346	346	/FTID=VAR_018935.
FT	VARIANT	493	493	E -> K.
FT	VARIANT	622	622	/FTID=VAR_018936.
FT	VARIANT	988	988	H -> Y.
FT	VARIANT	1019	1019	/FTID=VAR_018937.
FT	VARIANT	1019	1019	T -> M.
FT	VARIANT	1019	1019	/FTID=VAR_018938.
FT	VARIANT	1121	1121	V -> M.
FT	VARIANT	1121	1121	/FTID=VAR_018939.
FT	VARIANT	1121	1121	I -> M.
FT	CONFLICT	2	2	/FTID=VAR_018940.
FT	CONFLICT	120	120	R -> N (in Ref. 2).
FT	CONFLICT	175	175	R -> L (in Ref. 2).
FT	CONFLICT	753	753	R -> W (in Ref. 2).
FT	CONFLICT	787	787	Q -> R (in Ref. 2).
FT	CONFLICT	833	833	N -> S (in Ref. 2).
FT	CONFLICT	837	837	P -> L (in Ref. 3).
FT	CONFLICT	846	846	S -> G (in Ref. 2).
FT	CONFLICT	897	897	D -> E (in Ref. 2).
FT	CONFLICT	897	897	L -> F (in Ref. 3).
SQ	SEQUENCE	1203 AA:	130543 MW:	4FAB06999782C031 CRC64;
Query Match 99.8%; Score 4258; DB 1; Length 1203;				
Best Local Similarity 99.9%; Pred. No. 0;				
Matches 1202; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY	1	MTSPPLRELPPPTTPPARTAA	COLLAGS	KAPLMRAYFOGLIFSLGCGIQRHCGKVLF 60
DB	1	MTSPPLRELPPPTTPPARTAA	COLLAGS	KAPLMRAYFOGLIFSLGCGIQRHCGKVLF 60
QY	61	LGLLAFGALALGLRW	ETINLEQL	VEVGSRSQELHYTKEKLGEEAAYTSQMLQTAR 120
DB	61	LGLLAFGALALGLRW	ETINLEQL	VEVGSRSQELHYTKEKLGEEAAYTSQMLQTAR 120
QY	121	QEGENILTPKALCHLQAL	TASKVQVSL	YCKGSDLNKICYKSGVPLIENGMIENMEKL 180
DB	121	QEGENILTPKALCHLQAL	TASKVQVSL	YCKGSDLNKICYKSGVPLIENGMIENMEKL 180
QY	181	PCVILITPLDCTWEGAKLOGGSAY	PCRPDIQNTNLDPEQLLEELGPFASLEGFRELIDK 240	
DB	181	PCVILITPLDCTWEGAKLOGGSAY	PCRPDIQNTNLDPEQLLEELGPFASLEGFRELIDK 240	
QY	241	AQVGQAYVGRPCLDHDDLLHC	PPSAPNHS	RQAPNAVELSGSGCHGFSHKFMHQEELLG 300
DB	241	AQVGQAYVGRPCLDHDDLLHC	PPSAPNHS	RQAPNAVELSGSGCHGFSHKFMHQEELLG 300
QY	301	GMARDPOGELLRAEALOST	ELLASPROLTH	FRGQYQTHDIGNSEBOASTVLQAWORRV 360
DB	301	GMARDPOGELLRAEALOST	ELLASPROLTH	FRGQYQTHDIGNSEBOASTVLQAWORRV 360
QY	361	QLAQEALPENASQQIHAF	SPSTLDDILHAF	SEVSAARVVGGYLLMLAYACVTMLRWDCAQ 420
DB	361	QLAQEALPENASQQIHAF	SPSTLDDILHAF	SEVSAARVVGGYLLMLAYACVTMLRWDCAQ 420
QY	421	SOQSVGLAGVLVALAVAS	GLCALGITENAA	TOTLPLFALGIGVDDVFLAHAFTE 480
DB	421	SOQSVGLAGVLVALAVAS	GLCALGITENAA	TOTLPLFALGIGVDDVFLAHAFTE 480
QY	481	ALPGTPLQRMGECLORTGT	TSVLTISINNA	APFLMAALVPTALRAFSLQAATVVGCTVF 540
DB	481	ALPGTPLQRMGECLORTGT	TSVLTISINNA	APFLMAALVPTALRAFSLQAATVVGCTVF 540
QY	541	AVMLVPPAILSLDLRRRH	QCRDLVLC	CFSSPSCSAOVIQLPQELGGTVPGIAHATV 600
DB	541	AVMLVPPAILSLDLRRRH	QCRDLVLC	CFSSPSCSAOVIQLPQELGGTVPGIAHATV 600
QY	601	QAPTHCEASSQHWITL	PPQNLVPPSP	DPLASELPSFGSGSTRODLQCEENSTQKAAKCS 660

601	QAPTHCEASSQHVVITLPPQAHVLPPESSDPLGSELFSPGGSTRDLGLGGEETROKAAACKS	660
661	LPCARNLNLAHFARYQAFPLLQSHAKAIVLVLFAGALLGLSLYGATLVODGLALTVDPVPRG	720
661	LPCARNLNLAHFARYQAFPLLQSHAKAIVLVLFAGALLGLSLYGATLVODGLALTVDPVPRG	720
721	TTCNHLFLSNQLRYSLEVALVTQGCFDVYHSORALFDLKHQFESSLKAVLPPPTAQAPT	780
721	TREHAFDENQLRYSLEVALVTQGCFDVYHSORALFDLKHQFESSLKAVLPPPTAQAPT	780
781	WLHYRYNNWLOIAQAAPDQDWASGRITRHYSYRNGSEBDGALAYKLLOTGDAQEPDPSOLT	840
781	WLHYRYNNWLOIAQAAPDQDWASGRITRHYSYRNGSEBDGALAYKLLOTGDAQEPDPSOLT	840
841	TRKLVDRERGLIPPELFTMYCLTVMVSSDDPLGLMAASQANFYPPPPEMLHKDYDTTGENLRIP	900
841	TRKLVDRERGLIPPELFTMYCLTVMVSSDDPLGLMAASQANFYPPPPEMLHKDYDTTGENLRIP	900
901	PAQPLEPAQFPFLRLGRIGLQKTADFVEALEGARACACAEACGQGVHAYPGSGPFLFWEOVLGL	960
901	PAQPLEPAQFPFLRLGRIGLQKTADFVEALEGARACACAEACGQGVHAYPGSGPFLFWEOVLGL	960
961	RRCFFLLAVCILLCVCTFVCALLLINPWTAGLIVLYNMVTELFQIMGFGLIKLSAIPVV	1020
961	RRCFFLLAVCILLCVCTFVCALLLINPWTAGLIVLYNMVTELFQIMGFGLIKLSAIPVV	1020
1021	ILVASVGIGVEFTVHVVAALGFLTITQGSRNLRARAAHALEHTFAPVTDCAISTLLGLMLAGSH	1080
1021	ILVASVGIGVEFTVHVVAALGFLTITQGSRNLRARAAHALEHTFAPVTDCAISTLLGLMLAGSH	1080
1081	PDFIVRYFFFAALTVLTLGLLGHGLVLLPVLLSTLGCPPEVIOMVKESPESILSPAPOGCG	1140
1081	PDFIVRYFFFAALTVLTLGLLGHGLVLLPVLLSTLGCPPEVIOMVKESPESILSPAPOGCG	1140
1141	LRWGASSLSPOSFARVTTMTVAIHPPPLPGYAIHPADPEPWPSPAATSGCNLSSRGPGP	1200
1141	LRWGASSLSPOSFARVTTMTVAIHPPPLPGYAIHPADPEPWPSPAATSGCNLSSRGPGP	1200
1201	ATG 1203	
1201	ATG 1203	

RESULT 5
PTC2_MOUSE STANDARD; PRT: 1182 AA.

ID	PTC2_MOUSE	STANDARD;	PRT: 1182 AA.
AC	O35595;		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DE	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Patched protein homolog 2 (PTC2).		
GN	Name=Ptc2;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Embryo;		
RC	MEDLINE=98122566; PubMed=9462734;		
RA	Motoyama J., Takabatake T., Takeshima K., Hui C.-C.;		
RA	"Ptc2, a second mouse Patched gene is co-expressed with Sonic hedgehog.";		
RT	Nat. Genet. 18:104-106(1998).		
RL	[2]		
RN	SEQUENCE OF 196-446 FROM N.A.		
RC	STRAIN=BALB/c; TISSUE=Neuroretina;		
RX	MEDLINE=97379366; PubMed=9237688;		
RA	Takabatake T., Ogawa M., Takahashi T.C., Mizuno M., Okamoto M.,		
RA	Takeshima K.;		
RT	"Hedgehog and patched gene expression in adult ocular tissues.";		
RT	FEBIS Lett. 410:485-489(1997).		
CC	-!- FUNCTION: May have a role in epidermal development. May act as a		

THIS PAGE BLANK (USPTO)

CC receptor for Sonic hedgehog (SHH).
 CC SUBCELLULAR LOCATION: Integral membrane protein.
 CC TISSUE SPECIFICITY: Expressed in epithelial cells of the
 CC developing hair, tooth and whisker.
 CC DEVELOPMENTAL STAGE: Detected in 8.5 to 17.5 dpc embryos.
 CC SIMILARITY: Belongs to the patched family.
 CC SIMILARITY: Contains 1 sterol-sensing (SSD) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>
 CC or send an email to license@ebi.ac.uk).
 CC
 CC EMBL, AB010833; BAA24691.1; -
 CC F1R; T13952; T13952.
 CC MGD; MG1:1095405; Pch2.
 CC InterPro; IPR003392; Patched.
 CC InterPro; IPR004766; Patchedtm_recept.
 CC InterPro; IPR000731; SSD 5TM.
 CC Pfam; PF02460; Patched; 1
 CC TIGRfam; TIGR00918; 2A060602; 1.
 CC PROSITE; PSS0156; SSD; 1.
 CC Glycoprotein; Receptor; Transmembrane.
 CC DOMAIN 1 57 Cytoplasmic (Potential).
 CC TRANSMEM 58 78 Potential.
 CC DOMAIN 79 394 Extracellular (Potential).
 CC TRANSMEM 395 414 Potential.
 CC DOMAIN 415 428 Cytoplasmic (Potential).
 CC TRANSMEM 429 449 Potential.
 CC DOMAIN 450 457 Extracellular (Potential).
 CC TRANSMEM 458 478 Potential.
 CC DOMAIN 479 501 Cytoplasmic (Potential).
 CC TRANSMEM 502 522 Potential.
 CC DOMAIN 523 531 Extracellular (Potential).
 CC TRANSMEM 532 552 Potential.
 CC DOMAIN 553 686 Cytoplasmic (Potential).
 CC TRANSMEM 687 707 Potential.
 CC DOMAIN 708 963 Extracellular (Potential).
 CC TRANSMEM 964 984 Potential.
 CC DOMAIN 985 991 Cytoplasmic (Potential).
 CC TRANSMEM 992 1012 Potential.
 CC DOMAIN 1013 1013 Extracellular (Potential).
 CC TRANSMEM 1014 1034 Potential.
 CC DOMAIN 1035 1064 Cytoplasmic (Potential).
 CC TRANSMEM 1065 1085 Potential.
 CC DOMAIN 1086 1086 Extracellular (Potential).
 CC TRANSMEM 1087 1107 Potential.
 CC DOMAIN 1108 1182 Cytoplasmic (Potential).
 CC TRANSMEM 1182 1182 SSD.
 CC CARBOHYD 394 552 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 370 370 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 812 812 N-linked (GlcNAc...) (Potential).
 CC SEQUENCE 1182 AA; 129585 MW; 715233D912C32F2 CRC64;
 Query Match 89.3%; Score 5599; DB 1; Length 1182;
 Best Local Similarity 90.9%; Pred. No. 0;
 Matches 1074; Conservative 43; Mismatches 64; Indels 0; Gaps 0;
 QY 1 MTRSPRLRLPSPYTPPTAPQTLAGSLKAPLWLRAYFQGLFSLGCGIQHCKGVLF 60
 DB 1 MVRPLSLGELPSPYTPPTAPQTLAGSLKAPLWLRAYFQGLFSLGCGIQHCKGVLF 60
 QY 61 LGLAFALGALGRVAIETNLEQLWVGSRVSRQELHYTKELGEAAVTSOMLIQTAR 120
 DB 61 LGLVAFALGALGRVAIETNLEQLWVGSRVSRQELHYTKELGEAAVTSOMLIQTAR 120
 QY 121 QEGENTLTPEALGLHQAALTASKVQVSLYKSKNDLNKICYKGVPLIENGMIEMIKL 180
 DB 121 QEGGVLTPEALDLHQAALTASKVQVSLYKSKNDLNKICYKGVPLIENGMIEMIKL 180
 QY 181 PPCVILTPDLCFWEQAKLOGGSAYLPGRPDIQWNLDPQLLELGFASLEGFRELLOK 240

DB 181 FPCVILTPDLCFWEQAKLOGGSAYLPGRPDIQWNLDPQLLELGFASLEGFRELLOK 240
 QY 241 AQVGQYVGRPCPLHDDDLHCPSPAPNHSROAPNVAHELSCGCHGFKPKPMHQBELLG 300
 DB 241 AQVGQYVGRPCPLHDDDLHCPSPAPNHSROAPNVAHELSCGCHGFKPKPMHQBELLG 300
 QY 301 GMARDPQGLLRAEALQSTFLLMSPRQYEHFRGQYQTHDLCWSEEOASTVLAQWQRRFV 360
 DB 301 GTARDLQGLLRAEALQSTFLLMSPRQYEHFRGQYQTHDLCWSEEOASTVLAQWQRRFV 360
 QY 361 QLAQALPENASCOIHAHSSTLDDILHAFSEVSNARVGGYLLMLAYCATWMLWDCAQ 420
 DB 361 QLAQALPENASCOIHAHSSTLDDILHAFSEVSNARVGGYLLMLAYCATWMLWDCAQ 420
 QY 421 SOGSVGLAGVLLVALAVASGLGCLALLGITFNAATQVLPALALGIGVDVDFLLAHAFTE 480
 DB 421 SOGAVGLAGVLLVALAVASGLGCLALLGITFNAATQVLPALALGIGVDVDFLLAHAFTE 480
 QY 481 ALPCTPLQRMGECIQRCTGTSVLTSTINMAAFIIMAAALVPIPALRAFSLQAAIVVGTTFV 540
 DB 481 APDTPLPERMGECLSTGTSVLTSTINMAAFIIMAAALVPIPALRAFSLQAAIVVGTTFV 540
 QY 541 AVMLVFPAILSLDLRRHRCORLDVLCFSPSCSQAQVTOILPOELGDTGTVPVGIAHLTATV 600
 DB 541 AVMLVFPAILSLDLRRHRCORLDVLCFSPSCSQAQVTOILPOELGDTGTVPVGIAHLTATV 600
 QY 601 QAFTHCEASSQHVVTILPQAHVPPSPDPLGSELFSFGSTRDLGQEBETBQKAAKCS 660
 DB 601 QAFTHCEASSQHVVTILPQAHVPPSPDPLGSELFSFGSTRDLGQEBETBQKAAKCS 660
 QY 661 LPCARWNLAFARYOPAPILLOSHAKAIVLFCALLGLSLYGATLVODGLATDVVPRG 720
 DB 661 LLCNHTLAHARYOPAPILLOSHAKAIVLFCALLGLSLYGATLVODGLATDVVPRG 720
 QY 721 TXEHAFSLQALRYPSLYEVALVTQGGFYAHQSORALFDLHORFSSSLKAVLPPATQAPRT 780
 DB 721 TXEHAFSLQALRYPSLYEVALVTQGGFYAHQSORALFDLHORFSSSLKAVLPPATQAPRT 780
 QY 781 WLHYRNWNLQGIQAADFDDQD 840
 DB 781 WLHYRNWNLQGIQAADFDDQ 840
 QY 841 TRKLVDRGLIPRELYMGLTVWSSDPLGLAASQANPPPPPEWLHDKYDTTGENLRIP 900
 DB 841 TRKLVDRGLIPRELYMGLTVWSSDPLGLAASQANPPPPPEWLHDKYDTTGENLRIP 900
 QY 901 PAQPLEFAQFPFLRLGLOKTDVFAIEGARAACAEAGQAGVHAYPSGSPFLFWEQVYGL 960
 DB 901 PAQPLEFAQFPFLRLGLOKTDVFAIEGARAACAEAGQAGVHAYPSGSPFLFWEQVYGL 960
 QY 961 RRCFLAVCILVCTFLVCAILLNPMTAGLILVLAAMTVLFGIMFGIKLSAIPV 1020
 DB 961 RRCFLAVCILVCTFLVCAILLNPMTAGLILVLAAMTVLFGIMFGIKLSAIPV 1020
 QY 1021 ILVASVIGVETVHVALGFLTTCGSRNLRAHALEHTFAPVTDGALSTLGLLMLAGSH 1080
 DB 1021 ILVASVIGVETVHVALGFLTTCGSRNLRAHALEHTFAPVTDGALSTLGLLMLAGSH 1080
 QY 1081 FDIIVRYFAALTVTLGLLHGLVLLPVLISILGPPPEVIQMYKESPELSPAPQGGG 1140
 DB 1081 FDIIVRYFAALTVTLGLLHGLVLLPVLISILGPPPEVIQMYKESPELSPAPQGGG 1140
 QY 1141 LRWASSSSQPSQAPVTTMTVAHPPLPGAYTHAPDEP 1181
 DB 1141 LRWDREPTLPQSPARTVTTMTVAHPPLPGAYTHAPDEP 1181

RESULT 6
 CAC88120
 ID CAC88120 PRT: 1182 AA.
 AC CAC88120
 DT 02-MAR-2004 (T=EMBLrel. 27, C=created)

THIS PAGE BLANK (USPTO)